Problem 1

Problem 2

Problem 3

$$\frac{\partial}{\partial x}(\tanh(x)) = \frac{\partial}{\partial x}(\frac{e^x - e^{-x}}{e^x + e^{-x}})$$
$$= \frac{(e^x + e^{-x})(e^x + e^{-x}) - (e^x - e^{-x})(e^x - e^{-x})}{(e^x + e^{-x})^2}$$
$$= 1 - \tanh^2(x)$$

It substantially simplifies the calculations during backpropagation when computing the local gradient at an intermediate node and when passing the gradient to the lower layers.

Problem 4

Problem 5

Problem 6

Problem 7

Problem 8

Problem 9